ZANCO Journal of Pure and Applied Sciences

TYPE: Origin article
DOI http://dx.doi.org/10.21271/ZJPAS.38.1.9

OPEN ACCESS

*Corresponding author

Parzhin S. Hawez
parzhin.hawez@su.edu.krd

RECEIVED :19 /03/2025
ACCEPTED :13/08/ 2025
PUBLISHED :28/ 02/ 2026

KEYWORDS:

NHPP, Intensity function, Log-Linear Function, MLE, MoM, PSO, FFA, Simulation.

Parameter Estimation for a Log-Linear Nonhomogeneous Poisson Process Via Classical and Intelligent Approaches: A Case Study of Erbil International Airport

Parzhin S. Hawez and Khwazbeen S. Fatah

Department of Mathematics, college of science Salahaddin University-Erbil, Kurdistan Region, Iraq

ABSTRACT

The Nonhomogeneous Poisson Process (NHPP), characterized by rate functions that vary over time, is generally used to model the occurrence of events over time. A key challenge lies in accurately estimating the intensity function that derive the process. This paper focuses on estimating the parameters of a log-linear intensity function using the traditional Maximum Likelihood Estimation (MLE), **Method of Moments (MoM)** and two intelligent optimization techniques - Particle Swarm Optimization (PSO) and Firefly Algorithm (FFA). The objective is to evaluate and compare these estimation methods to identify the most effective approach, followed by the application of a NHPP to model the number of passenger arrivals at Erbil International Airport (EIA) - an important factor for improving airport facility planning. The analysis utilizes passenger arrival data recorded at Erbil International Airport (EIA) in Kurdistan Region (KR) of Iraq from January 1, 2021, to September 30, 2024. The passenger arrival rate is initially analyzed through a homogeneity test, and then the Cross-Validation technique is applied to evaluate the model's predictive performance. Additionally, two simulation techniques are employed to assess the performance of the estimation methods by calculating the Root Mean Square Error (RMSE). The findings reveal that the PSO technique outperforms MLE, MoM and FFA, providing more accurate parameter estimates and faster convergence for the log-linear model.

1.Introduction

A counting process is a type of stochastic process characterized by values that are always not declining and not-negative. Among the most prevalent counting procedure is the Poisson process that models the number of events occurring over a given time for various real-world phenomena. In a Homogeneous Poisson Process (HPP) defined by a fixed rate parameter λ, the time intervals between occurrences are independent and adhere to an exponential distribution. Nonetheless, this presumption of a constant rate may not be applicable in real-world situations where events happen at fluctuating rates throughout time. To speech this issue, a NHPP has been proposed as a more generalized form of the HPP.

The NHPP model is particularly useful for analyzing events with time-varying occurrence frequencies, as defined by its intensity function. Its broad applicability across diverse scenarios is mainly due to the inherent independence property of events in the model. Examples include applications such as analyzing the number of accidents, assessing software reliability and improvement, and studying phenomena in fields like astronomy, statistical mechanics, telecommunications, biology, economics, and geology. Examples of NHPP traffic accidents,(Frenkel et al., 2003, Sugiyanto and Santi, 2017, Zhao et al., 2021, Hussein and Hassan, 2023, Chen and Lü, 2024) and various other domains. Several approaches had been suggested to estimate the parameters of the intensity function in NHPP; they enclose parametric methodologies (Xie et al., 1996, Ofori-Addo, 2023, PARK, 2024), Bayesian methods and non-parametric approaches, (Chang et al., 1989, Li et al., 2020, Aminzadeh, 2025). The loglinear NHPP has mostly been used to represent systems in transportation - including air transportation - and telecommunications (Frenkel et al., 2003, Yarmohammadi et al., 2022, SH et al., 2023).

Iraq, a federal developing nation, offers a range of transportation options, particularly in land and air travel. The national transportation system plays a vital role in promoting overall development, with air transportation serving as a key component that supports economic growth across all regions of the country. To gain a comprehensive understanding of the air transport sector, multiple indicators must be considered. A major goal in transportation development is to establish an efficient air transportation system, where understanding passenger arrival patterns is essential for effective airport terminal planning. The KR of Iraq, an autonomous area within the federal structure, is entering the new century with significant progress. Known for its rich cultural heritage, historical depth, and abundant resources, KR has seen notable advancements in air transport—most prominently represented by EIA (Ghareeb, 2024; Shareef, 2025; Abdulrazaq, 2025). Despite these developments, there is still a notable shortage of research on air transportation in Iraq, including the KR. This gap exists even though air transport is essential for national development and plays a key role in promoting economic growth throughout the country.

This research proposes modeling passenger arrival rate at EIA for the period from January 1, 2021, to September 30, 2024, using a NHPP. A log-linear function is used to model the passenger arrival rate due to its characteristics that ensure independence, non-negativity, and a consistently non-decreasing pattern. Parameter estimation for the log-linear rate function is performed using classical approaches including MLE and MoM alongside two metaheuristic algorithms: PSO and FFA. To evaluate the effectiveness of these methods, two simulation techniques are applied.

This research is prepared as shadows: Section 2 introduces a summary of the key impressions related to the NHPP. Using a statistical hypothesis test to ascertain if the arrival rate is homogeneous or inhomogeneous. Additionally, a technique for modeling the NHPP's intensity function is provided. In section 3, an application is provided and an NHPP to model the passenger arrivals with log-linear rate function is proposed. A simulation study is performed in section 4 with result discussion. In section 5 the main conclusion is given.

2. Poisson Process

A Poisson process is a distinct kind of

counting process utilized to simulate the occurrence of events that transpire independently and at a constant average rate. as stated by (Ross, 2014), the following definition related to Poisson Processes are provided below:

2.1 . Counting process

$\{N(t), t \geq 0\}$ is a stochastic counting process that meets the following criteria:

1) Initial Condition: $N(0) = 0$
2) Non-Negativity: $N(t) \geq 0$ for all t ≥ 0;
3) Integer valued: $N(t)$ takes only integer values;
4) Monotonicity: $N(t)$ is non-decreasing, meaning that if $s < t$ then $N(s) \leq N(t)$;
5) **Independent Increments:** The quantity of events happening in disjoint time intervals is independent
6) **Stationary Increments:** the distribution of the number of Actions within any time span are contingent only upon the duration. of the interval, not on its starting point for HPP.

2.2 Homogeneous Poisson process (HPP)

A counting process $\{N(t), t \geq 0\}$ is referred to as an HPP with the rate λ (λ>0) if the subsequent criteria are met:

1 $N(0) = 0$;
2 $N(t)$ has independent increments;
3 $P(N(t+h) - N(t) = 1) = \lambda h + o(h)$ and
4 $P(N(t+h) - N(t) \geq 2) = o(h)$, where $o(h)$ is a tiny amount applied in the condition and $h > 0$. $\lim_{h\to 0} \frac{o(h)}{h} = 0$. (Ross, 2014)

In accordance with the conditions described above, events within the time interval t are distributed as follows the Poisson distribution with the rate λt. If $N(t)$ is a Poisson process with rate $\lambda(\lambda > 0)$, then for every $t, s > 0, N(t+s) - N(s)$ is a Poisson random variable with a mean λt.The HPP is based on the assumption of a constant occurrence rate. However, this assumption may not be valid for most natural phenomena, making the HPP unsuitable for such cases. Instead of treating as a random variable the intensity function and estimating it through various strategies, a more straightforward approach is to use the NHPP.

2.3Non-Homogeneous Poisson process (NHPP)

The process $\{N(t), t \geq 0\}$ is an NHPP with an intensity function λ(t) if the following necessities are met:

1 $N(0) = 0$.
2 $N(t)$ are independent increments.
3 $P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$ and
4 $P(N(t+h) - N(t) \geq 2) = o(h)$, where h > 0 and o(h) represents a negligible quantity that meets the specified criterion that $\lim_{h\to 0} \frac{o(h)}{h} = 0$.

Using the intensity function λ(t), the mean function of the NHPP $\{N(t), t \geq 0\}$ is defined as:

$$m(t) = \int_0^t \lambda(t)\, dt \qquad (1)$$

Furthermore, based on Equation 1, the estimation of the mean average N(t) is equivalent to:

$$\hat{m}(t) = E\big(N(t)\big) = \int_0^t \hat{\lambda}(x)dx \qquad (2)$$

where $E(.)$ denotes the expected value. Under the aforementioned requirements, it is demonstrated that for any $t, s > 0, N(t+s) - N(s)$ follows a Poisson random variable with a mean defined as follows:

$$m(t+s) - m(s) = \int_s^{t+s} \lambda(x)\mathrm{d}x \qquad (3)$$

As a outcome, $\{N(t), t \geq 0\}$ has an NHPP according to the subsequent distribution:

$$P(N(t) = k) = \frac{\left(\int_0^t \lambda(x)\mathrm{d}x\right)^k}{k!} e^{-\int_0^t \lambda(x)\mathrm{d}x}\ k = 0,1,2,\ldots \qquad (4)$$

Before estimating the parameters of the intensity function, it is essential to explicitly evaluate the assumption of homogeneity. If the data prove to be inhomogeneous, applying an NHPP would be appropriate. Therefore, a test is implemented to determine whether the observations satisfy the homogeneity assumption, complemented by a graphical tool to assess whether the dataset conforms to a

specified theoretical distribution. In addition, Cross-validation, which is a statistical technique used to evaluate how well a predictive model, is performed.

2.4 Poisson Process Homogeneity Test:

Brown and Zhao (2002) introduced a novel homogeneity test that has shown superior performance compare to the likelihood ratio, conditional chi-square, and Neyman-Scott (DasGupta and DasGupta, 2008, Brown and Zhao, 2002). To assess the uniformity of passenger arrival rates at EIA, this test is utilized, assuming that $N_1(t), N_2(t), \ldots, N_n(t)$ are independent non-negative integer-valued random variables tests. For this test, the following hypothesizes are considered:

$$H_0: N_i \sim \text{Poiss}\,(\lambda_i),\ \lambda_1 = \lambda_2 = \cdots = \lambda_n$$

$$H_1: N_i \sim \text{Poiss}\,(\lambda_i),\ \sum_{i=1}^{n} \left(\lambda_i - \bar{\lambda}\right)^2 > 0$$

This explains that if $N \sim \text{Poiss}\,(\lambda)$, then:

$\text{Var}_\lambda\left(\sqrt{N+\frac{3}{8}}\right) = \frac{1}{4} + o\left(\frac{1}{\lambda}\right)$, where $Y_i = \sqrt{N_i + \frac{3}{8}}$. The test statistic T can be defined as:

$$T = 4\sum\ (Y_i - \bar{Y})^2 \ \sim\ \chi_{(n-1)}(4\sum\ (\nu(\lambda_i) - \bar{\nu}_n)^2) \qquad (5)$$

in which, $\bar{\nu}_n = \frac{1}{n}\sum\ \nu(\lambda_i), \nu(\lambda_i) = E_{\lambda_i}(Y_i) = E_{\lambda_i}\left(\sqrt{N_i + \frac{3}{8}}\right)$

H_0 is tested against H_1; H_0 is rejected if $T > \chi^2_{(n-1),(1-\alpha)}$.

2.5 Quantile-Quantile (Q-Q) Plot

A graphical technique for determining if a dataset adheres to a specific theoretical distribution is the Q-Q plot., such as the normal or exponential distribution. It works by plotting the quantiles of the observed data against the quantiles of a chosen reference distribution. Quantiles divide the data into intervals with equal probabilities. If the data closely follow the reference distribution, the plotted points will fall around along a straight diagonal line. The points' proximity to this line, the more likely it is that the dataset matches the theoretical distribution (Ismael, 2024, Bening et al., 2025).

2.6 Cross-validation

Cross Validation is a statistical technique used to evaluate how well a predictive model can perform. It involves dividing the available dataset into subsets or "folds", training the model on some of these subsets, and testing the model on the remaining complementary subsets. This process is repeated multiple times with different data partitions to ensure a reliable and unbiased estimate of how well the model will perform on unseen data. This technique ensures that each data points has an equal probability of appearing in the training and test sets, offering a reliable assessment of the model's performance (Jiang and Simon, 2007, Senceroglu et al., 2022, Ali and Al-Zubaidi, 2025).

The primary goal of cross-validation is to prevent overfitting. In other words, the model works well with training data but not with data that is new. K-fold cross-validation is a popular cross-validation technique. K-fold cross-validation is frequently suggested as a solution to the challenge of achieving accurate performance estimation and classification. It divides the data into k components of equal sizes. After training the model k times, the final 1-fold was tested (Oyedele, 2023, Patel et al., 2024). This method is one of the recommended approaches for validating machine learning models since it guarantees that the model's performance is assessed on several data splits. It can learn more about how the model works with unknown data by employing cross-validation, which will lead to more precise and trustworthy predictions.

2.7 Parameter Estimation

The rate or intensity function plays an important role in describing the probabilistic behavior of the NHPP. Common choices for modeling this function include trigonometric, polynomial, exponential, and spline forms. For non-negative counting data, the exponential or

log-linear function is frequently selected since it can guarantee non-negativity. Parameter estimation for NHPPs has been the focus of extensive research studies and generally falls into two main methodological categories. Analytical approaches such as MLE and LSM are often preferred due to their high accuracy (Song and Gu, 2004, SH et al., 2023). Numerical techniques - including the Newton-Raphson method and Monte Carlo simulation, are often applied to address the complexity of models that arise from analytical methods like MLE, MoM, and LSM. In recent years, artificial intelligence-based techniques - such as PSO, FFA, Gray Wolf Optimizer (GWO) among others – have attracted significant attention (Kilinc and Yurtsever, 2022, Abdel-Basset et al., 2024). These approaches have been rapidly adopted in various optimization problems and are now widely applied to address complex parameter estimation problems.

In this research, a log-linear function is used to represent the passenger arrival rate λ(t), as it ensures independence, non-negativity, and a non-decreasing pattern (Frenkel et al., 2003, Wang et al., 2018, Yarmohammadi et al., 2022). accordingly, based on the properties discussed earlier, the arrival rate function λ(t) is modeled using a log-linear function defined as:

$$\lambda(t) = e^{\alpha+\beta t}, \quad where \ \ \alpha, \beta > 0 \ and \ t > 0 \qquad (6)$$

2.7.1 Maximum Likelihood Method

Suppose that a NHPP with intensity function $\lambda(t) = e^{\alpha+\beta t}$ is observed over a fixed interval $(0, T)$. Let n denotes the total quantity of events transpiring during this time frame. If these events occur at $0 < t_1 < t_2 < \cdots < t_n \leq T$, then the likelihood function is given by;

$$L(\alpha,\beta) = \left(\prod_{i=1}^{n} \lambda(t_i)\right) \exp\left(-\int_0^T \lambda(s)ds\right)$$

$$= \exp\left(n\alpha + \beta\sum_{i=1}^{n} t_i - \frac{exp(\alpha)}{\beta}(exp(\beta T) - 1)\right) \qquad (7)$$

Then, the log-likelihood function, will be given by;

$$l(\alpha,\beta) = \log(L)$$

$$= n\alpha + \beta\sum_{i=1}^{n}\frac{e^{\alpha}\left(e^{\beta T}-1\right)}{\beta} \qquad (8)$$

The maximum likelihood estimators $\hat{\alpha}$ and $\hat{\beta}$ are the respective values of α and β that maximize $l(\alpha,\beta)$. By executing partial differentiation of (3) with regard to α and β, we get the score functions which when equated to zero yields the following system of equations:

$$e^{\hat{\alpha}} = \frac{n\hat{\beta}}{e^{\hat{\beta}T}-1} \qquad (9)$$

$$\sum_{i=1}^{n} t_i = \frac{nT}{1-e^{-\hat{\beta}T}} - \frac{n}{\hat{\beta}} \qquad (10)$$

Equations (9) and (10) are solved simultaneously to obtain the maximum likelihood estimators for α and β. Since explicit solution for equation (10) does not exist, it may be solved numerically using the Newton Raphson method (Veronica et al., 2014).

2.7.2 Method of Moments (MoM)

The MoM is a classical and straightforward technique for estimating the parameters of a probability distribution or statistical model. It was introduced by Karl Pearson in the late 19th century. The core idea of MoM is to equate the theoretical moments of a distribution (calculated based on unknown parameters) with the sample moments (calculated from observed data), and then solve the resulting equations to estimate the unknown parameters (Murray and Bellhouse, 2020, Ghilli et al., 2024). In statistics, a moment is a quantitative measure related to the shape of a distribution. The definition of a random variable X's $k-th$ theoretical moment is:

$$\mu_k' = E[X^k]$$

is the $k-th$ sample moment is calculated as:

$$m'_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

The first moment represents the mean, the second central moment represents the variance, and so on. When MoM is applied to estimate parameters for the log-linear NHPP with $\lambda(t)$ defined as:

$$\lambda(t) = \exp(\alpha + \beta t)$$

The cumulative intensity function (mean function) over the interval $[0, T]$ is:

$$\Lambda(T) = \int_0^T \lambda(t)dt = \frac{e^{\alpha}}{\beta}\left(e^{\beta T} - 1\right)$$

When MoM is applied, $\Lambda(T)$ is equated to the observed number of events n :

$$n = \frac{e^{\alpha}}{\beta}\left(e^{\beta T} - 1\right) \quad (11)$$

For a second moment condition, the sample mean of arrival times $\bar{t}$ is calculated:

$$\bar{t} = \frac{1}{\Lambda(T)}\int_0^T t\,\lambda(t)dt \quad (12)$$

These equations, 11 and 12, form a nonlinear system that can be solved numerically to obtain the estimates α and β.

2.7.3 Particle Swarm Optimization Algorithm (PSO):

PSO technique, which was initially introduced by Eberhart and Kennedy in 1995 (Bratton and Kennedy, 2007), is a relatively recent technique for addressing various optimization problems and has recently gained recognition as an effective method for parameter estimation problems; The particle location and velocity are repeatedly updated in this population-based stochastic optimization approach. PSO is derived from the cognitive and social dynamics shown by some animal groups, such as avian flocks or piscine schools (SH et al., 2023, Hassen and Abdlrazaq, 2024).

The PSO method sustains a population of particles, each representing a possible solution to an optimization problem, with respectively result including a set of parameters in multidimensional space. The particle's location represents a viable, if not optimal, solution to the challenge.

Optimal advancement is necessary to adjust the particle's location to enhance the goal function's value. Subsequently, the fundamental principle of the PSO algorithm is that a population (swarm) of particles is perpetually relocating within a search space in order to identify a solution. Each particle, denoted by i, is represented by its position (x_i) and velocity (v_i), as determined by a straightforward formula. The PSO algorithm can be explained by the following steps:

- Establish a random initial location by giving each particle a random position and velocity;
- Calculate the fitness value for each particle.
- Identify the location of the optimal fitness value, the local best (p-best), for each particle and update it if it surpasses the previous value.
- Determine the location of the optimal fitness value among all particles, referred to as the global best (g-best), and update it if it surpasses the previous value.
- Execute all steps until the termination requirements are met.

The following equations are used to modify the location and velocity of each particle:

$$V_i(t+1) = \omega V_i(t) + C_1 r_1(P_i - X_i(t)) + C_2 r_2(G - X_i(t))$$

$$X_i(t+1) = X_i(t) + V_i(t+1)$$

where the velocity vector of particle s in time t is indicated by $V_i(t)$; The position vector particles in t time are described by $X_i(t)$;The particle's personal best position is denoted by P_i. ; G is the best position of the particle creating at present; ω denotes initial weight; C_1, C_2 are two acceleration numbers, called cognitive and social parameters, correspondingly; Two random functions in the range are r_1 and r_2.

2.7.4. Firefly Algorithm (FFA)

The FFA was first proposed by Xin-She Yang at Cambridge University(Arora and Singh, 2013), The methodology of this algorithm is predicated on the nocturnal flashing patterns and behaviors of fireflies. To attain optimal solutions for optimization problems, the objective function employs the differences in light intensity and the formulation of attraction. Both help the fireflies to move towards brighter and more attractive locations. To formulate the FFA, the flashing features may be perfect according to the subsequent tierce principles.(Fister et al., 2013):

1. All fireflies are unisex, allowing one firefly to be attracted to another irrespective of their sex.
2. The attraction of each firefly is directly proportional to its brightness, which reduces with increasing distance from another firefly. Consequently, for each pair of blinking firefly types, the dimmer specimen will advance towards the brighter one, and both will descend as the space between them expands. If one firefly is not more luminous than another, it moves erratically;
3. The luminosity of the firefly will be influenced or dictated by the value of the goal function to be optimized; the brightness may be directly proportional to the objective function. Alternative manifestations of brightness can likewise be delineated in a method akin to the fitness function utilized in genetic or bacterial foraging algorithms (BFA).

Consequently, according to the second rule, a specific firefly will be drawn to those that exhibit greater brightness, and the allure of a firefly can be quantified by its luminosity or light intensity, which is intrinsically linked to the value of the objective function. In the most straightforward scenario for maximum optimization problems, the light intensity $I(x)$ of a firefly at a specific position x is proportional to the value of the fitness function, $I(x) \propto f(x)$. The allure of fireflies, represented by β, is directly proportional to their luminescence intensity. $I(r)$ should therefore vary with the distance rij between firefly i and firefly j. As light intensity diminishes with increasing distance from its source and is also absorbed by the medium,

The appeal would fluctuate according to the level of absorption. The light intensity $I(r)$ varies monotonically and exponentially with the distance r.(Kwiecień and Filipowicz, 2012):

$$\mathrm{I(r)} = \mathrm{I_0 e^{-\gamma r}}$$

Where I_0 denotes the initial light intensity and γ signifies the light absorption coefficient.

Elementary physics demonstrates that the intensity of light is inversely proportional to the square of the distance, denoted as r, from the source. Consequently, the relationship between attractiveness β and distance r will be defined as:

$$\beta r = \beta_0 e^{-\gamma r^2}$$

It is important to note that the exponent γr can be substituted with alternative functions such as γr^m when $m > 0$; hence, the FFA can be summarized as pseudo code. The distance between any two fireflies i and j, located at x_i and x_j, can be represented as the Cartesian distance.

$r_{ij} = \left\|\mathrm{x_i} - \mathrm{x_j}\right\|_2$ or the ℓ_2-norm. The motion of firefly i is drawn towards a more luminous firefly j, and the position update of firefly i at X_i will be as follows:

$$\mathrm{X_{i+1}} = \mathrm{X_i^t} + \beta_0 \mathrm{e^{-\gamma r_{ij}^2}}\left(\mathrm{X_j^t} - \mathrm{X_i^t}\right) + \alpha_t \in_i^t$$

In Equation 3, the second component represents the attraction to X_j, while the final term accounts for randomness, with α_t serving as a randomization parameter constrained between 0 and 1. The term $\in_i^t$ denotes a vector of random values sampled from a Gaussian, Uniform, or other distribution at time t.

In summary, FFA is ruled by three parameters: the randomization parameter α, the attractiveness β, and the absorption coefficient.(Hsu, 2022, SH et al., 2023, Szemis et al., 2024, Yu et al., 2025).

3.Application

In this section, to bridge the theoretical framework with practical application, an NHPP model is employed on a real dataset documenting arrival at EIA, the main airport in the KR, which has witnessed considerable improvements in air transport infrastructure (Ghareeb, 2024; Shareef, 2025; Abdulrazaq,

2025). Despite these developments, there is still a lack of research focused on air transportation in Iraq, including the KR. This gap remains even though air transport is essential for national development and plays a pivotal role in supporting economic growth throughout the country. Given the significant role of air transportation in driving economic activities, various indicators are required to comprehensively evaluate its current and future state. One key objective is to develop an efficient air transportation system, where the pattern of passenger arrivals is essential for planning airport improvements. This paper aims to model the passenger arrival rate at EIA from January 1, 2021, to September 30, 2024.

3.1 Data Collection:

The dataset reflecting the number of arrivals at EIA was obtained from the EIA administration office and covers a 45-month period, from January 1, 2021, to September 30, 2024. The data were analyzed, and the intensity—calculated as the monthly number of arrivals at EIA divided by the number of days in each month—was computed for every interval within this period.

Table 1 presents the arrival data across 45 months, while Table 2 summarizes the corresponding arrival intensities. The analysis reveals that the arrival rate is non-negative and shows an increasing trend over the observed period. Furthermore, Figures 1, 2 provide graphical illustrations of the arrival patterns at EIA during this timeframe.

3.2 Model Formulation for Passenger-Arrivals

Passenger arrivals at EIA airport can be modeled as an NHPP, where the number of arrivals in any given time interval is independent of other intervals and governed by a time-dependent intensity function, λ(t). This intensity function, which serves as the key parameter of the process, is characterized by a log-linear form as specified in equation (6). The log-linear model was selected based on an initial exploratory analysis of the dataset and its widespread use due to its desirable mathematical properties—most notably, ensuring non-negativity and providing the flexibility to accurately capture temporal variations in arrival rates (Bhaskar et al., 2020, Guo et al., 2022, Ghilli et al., 2024, Venkataraman and Rumpler, 2025).

In analyzing count data statistically, it is common practice to model the logarithm of $\lambda(t)$ rather than $\lambda(t)$ itself. This log transformation guarantees that intensity estimates remain non-negative. In this work, monthly arrival patterns are modeled $\log(\lambda(t)) = \alpha + \beta t$, where α, β are model parameters showed in Figure 3. Model validation is performed and parameter estimation is carried out using four different methods. This modeling framework enables accurate forecasting of arrival patterns, which is crucial for effective airport resource management.

3.3 Model Validation:

Validating an NHPP model is a necessary step to confirm that it accurately reflects the observed arrival process. To assess the suitability of an NHPP with a log-linear intensity function for modeling passenger arrival rates, the following techniques are applied:

3.3.1 Homogeneity Test for Passenger Arrival Rate:

To evaluate whether an NHPP with a log-linear intensity function is appropriate for modeling passenger arrival rate at EIA, a homogeneity test defined by a test statistic introduced in section 2.4 is applied. The test statistic, as defined by Equation 5, is then computed after transforming the data into: $Y_i = \sqrt{N_i + \frac{3}{8}}$; $T = 4\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = 301118.4772$, for $\alpha = 0.05$,

$$T = 301118.4772 > \chi^2_{(44;0.95)} = 61.65$$

Therefore, at a 0.05 significance level, the hypothesis of homogeneity for the total number of arriving passengers at EIA is rejected, indicating that an NHPP is a suitable model for the data under study.

3.3.2 Quantile-Quantile (Q-Q) Plot

The Q-Q, introduced in section 2.6, is a graphical technique used for model validation. A Q–Q plot was constructed to evaluate the goodness of fit for the proposed model. Event times were transformed using the estimated cumulative intensity function of the log-linear NHPP for the arrival dataset. The transformed

interarrival times, which under the model are expected to follow an Exponential (1) distribution, were compared to the theoretical quantiles. The plot showed that the observed quantiles closely followed the reference line, supporting the adequacy of the model.

Figure 4 shows the Q-Q plot between empirical Quantiles and Z-Score.

3.3.3 Cross Validation

In this study, we applied a cross-validation procedure by dividing the dataset into five parts, employing 5-fold cross-validation. We used four of these parts, generated through the MATLAB-based PSO software, to optimize the model parameters. The remaining fifth part was then evaluated, and its trend line was plotted. Further analysis showed that the calculated $R^2 = 0.97$, as seen in Figure 5. We therefore came to the conclusion that dataset representing arrivals at EIA has a log-linear function.

3.4 Parameter Estimation for the Proposed Model:

When the actual passenger arrival data at EIA is modeled using a log-linear function, parameter estimation is necessary to be carried out. Classical approaches such as MLE and MoM, along with intelligent optimization techniques, specifically PSO and FFA are used. Each of these methods is described in detail below.

3.4.1 Parameter Estimation using MLE:

To estimate the parameters of the future log-linear $\lambda(t)$ using the actual arrival data with MLE, as described in Section 2.7.1, the passenger arrivals were aggregated into non-overlapping time intervals. As a result, the total number of arrivals per interval was used instead of precise arrival times. Using the statistical software SPSS, the estimated parameters of the log-linear function are as follows: $\hat{a} = 7.246$ ، $\hat{b} = 0.001$ with the rate function, which represents the arrival rate at EIA from 1/1/2021 to 30/9/2024, is given by the following:

$$\hat{\lambda}(x) = \mathrm{e}^{7.246+0.001x}$$

Then, the mean-value function $\Lambda(t)$, which is defined by equation (2), is determined.

$$\begin{aligned}\Lambda(t) = \mathrm{E}\big(\mathrm{N(T)}\big) &= \int_0^t (\mathrm{e}^{7.246+0.001x})dx \\ &= 1000\mathrm{e}^{7.246+0.001\mathrm{t}} \\ &- 1000e^{7.246} \\ &= 1000e^{7.246}(e^{0.001t} - 1) \\ &= 1{,}402{,}480\ \ (e^{0.001t} - 1)\end{aligned} \quad (13)$$

Hence, the distribution function for number of arrivals at EIA is provided by;

$$P\{N(t) = k\} = \frac{(\Lambda(t))^k}{k!} e^{-\Lambda(t)}$$

$$= \frac{(1{,}402{,}480\ \ (e^{0.001t} - 1)\ )^k}{k!} e^{1{,}402{,}480\ \ (e^{0.001t}-1)} \quad (14)$$

With mean: $\Lambda(t) = 1{,}402{,}480\ \ (e^{0.001t} - 1)$, $where\ t \geq 0$.

Then, the quantity of arrivals at a given time period can be predicted using the interval's length s:

$$P\{N(t+s) - N(t) = k\}$$

$$= \frac{(\Lambda(t+s) - \Lambda(t))^k}{k!} e^{-(\Lambda(t+s)-\Lambda(t))} \quad (15)$$

and the expected value E($N(t+s) - N(t)$) is determined from equation (3).

For example: if it is required to predict the number of arrivals At EIA for the period October,1/10/2024 to 1/11/2024, for time interval [1371,1402), where the length of the interval is $s = 31$ and $t = 1371$. By using equation (3), the following is obtained.

$$\begin{aligned}\Delta(1371,31) &= E(N(1402) - N(1371)) \\ &= \int_{1371}^{1402} \mathrm{e}^{7.246+0.001t}\, dt = 173950.70\end{aligned}$$

$$\begin{aligned}\sigma(1371,31) &= \sigma(N(1402) - N(1371)) \\ &= \sqrt{\Delta(1371,31)} = \sqrt{173950.70} \\ &= 417.07\end{aligned}$$

This indicates that the predicted average number of arrivals at EIA between October 1, 2024, and November 1, 2024, is approximately 173950.70 passengers, with a standard deviation of 417.07.

3.4.2 Parameter Estimation using MoM:

To apply the MoM, as described in Section 2.7.2, for the proposed log-linear function fitted to the real datas utilized, and the arrival rate function was subsequen parameter estimates for the log-linear function were as f and $\hat{\lambda}(x) = \mathrm{e}^{7.20+0.0008x}$.

3.4.3 Parameter Estimation using PSO:

In this section, the PSO method, as described in Section 2.7.3, is applied to estimate the parameters of an NHPP with a log-linear rate function using the real arrival dataset. The estimation technique is detailed in the algorithm that follows.

Algorithm (1): Parameter Estimation of the log-linear function using PSO

Step1: Identify each of: the number of particles $(N = 45)$; the maximum iterations $(\max(i) = 100)$; the acceleration coefficients C_1 and C_2 , where $C_1 = C_2 = 1$; $r_1 \, and \, r_2$ are random numbers within the interval $[0,1]$. The inertial weight's minimum and maximum values are: $\max(w) = 0.9 \, and \, \min(w) = 0.4$.

Step 2: Specify randomly each particle's location, with each position serving as an estimate for the loglinear process parameters and each particle's beginning position being produced from a uniform distribution within the range $[\alpha, \beta\,]$ in [7,8] and [0.001,0.005] ;

Step3: Initialize each particle's starting velocity using a uniform distribution.

Step4: Define the fitness function as the Mean Squared Error (MSE), defined by:

$$MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{n}$$, where y_i represent parameters and $\hat{y}$ is the estimators

Step5: update the parameters estimator $\hat{\beta}, \hat{\alpha}$ for the process based on the resulting value of the MSE function by updating the speed (V_i) utilizing the equation that follows:

$$V_j^{(i)} = \gamma V_j^{(i-1)} + C_1 r_1 \left[P_{\text{best},j} - X_j^{(i-1)}\right] + C_2 r_2 \left[G_{best,j} - X_j^{(i-1)}\right] j = 1,2 \ldots, N$$

Subsequently, updating locations X_i in accordance with the equation:

$$X_j^{(i)} = X_j^{(i-1)} + V_j^{(i)}; j = 1,2, \ldots, N$$

Step6: Repeat steps $4 - 5$ until an $max(i)$ is reached.

To apply this algorithm to the dataset, a MATLAB program is written, by running this program the parameters estimate is obtained as: $\hat{a} = 7.832$, $\hat{b} = 0.0002, and \quad \hat{\lambda}(x) = e^{7.832+0.0002x}$

3.4.4 Parameter Estimation using (FFA).

To estimate the log-linear parameters using the FFA, as introduced in Section 2.7.4, Algorithm 2 presents the key steps involved in estimating the parameters of the log-linear process:

Algorithm (2): Parameters Estimation of the log-linear process using the FFA:

Step1: Establish the particle count at N=45 and the iteration count at $i_{\max} = 100$;

Step2: Initialize the input parameters: $\alpha_0 = 0.5$, $\beta_0 = 0.2$, $\gamma = 1$, $n = 2$, $\in = 5$;

Step3: Identify the position x_i of every particle i, with every position signifying an estimation for the parameters of the loglinear Process;

Step4: Specify the fitness function MSE, in which

$MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{n}$; where y_i represent parameters and $\hat{y}$ is the estimate

Step5: Execute the FFA algorithm utilizing the specified values:

$r = r_{ij} = \|x_i - x_j\|_2$; $\beta = \beta_0 * \exp(-\gamma * r^2)$; $steps = \alpha_0 * (rand(1,n) - 0.5) * \in$; $Xnew = Xi + \beta * (Xj - Xi) + steps$;

Step6: Compute each value based on the outcome value of the fitness function MSE.

Step7: repeat steps 4 - 5 until an $i_{\max}$ is attained.

When a written program in MATLAB is utilized to implement Algorithm 2, the obtained parameter estimates are:

$\hat{a} = 7.0$ ، $\hat{b} = 0.003$, and $\hat{\lambda}(x) = e^{7+0.003\,x}$

The primary results obtained from the four approaches - MLE, MoM, PSO, and FFA - are shown in Table 3 and Figure 6. The outcomes from these estimation methods show that the estimates for β are generally similar.

4.Simulation:

Simulation is a computer-based technique that enables testing on an accurate digital model. It can be used to imitate the operation of an existing or proposed system or a process and analyze its behavior by exploring different values of the key parameters involved. The methodology involves creating a model of the process under study, implementing it as a computer program, using a digital computer to run it, and then evaluating the results found. Additionally, simulation is among the most effective approaches for producing random

variables from specific probability distributions. In this section, two commonly used methods for simulating an NHPP are presented and then their algorithms are compared in terms of their efficiency to arrive at the optimal estimate for the log-linear parameters.

To simulate random variables from a log-linear function for the passenger arrivals at EIA and then estimate its parameters, the main steps are described below (SH et al., 2023):

4.1Virtual Value Assignment:

In this phase, considered as the foundation for the other stages, the necessary input values are provided to generate random numbers from a log-linear function. These values are as follows:

1) Sample Size (N): The sample sizes used in this paper are: N=15, N=30 and N=45;
2) Parameter values for log-linear process: The probability distribution function for log linear Process used to generate random variables assuming two values for each parameter: $\hat{\alpha} = 7\ ,\ 7.5\ ;\ \hat{\beta} = 0.001\ , 0.003;$
3) Sample Repetition Size (i): To obtain highly consistent results, each experiment was performed once with $i = 45$.

4.2 Simulation Methods:

To generate random variables from log-linear function two main techniques are used:

4.2.1 Time-Scale Transformation Method:

The Inverse Method (also known as the Time-Scale or Transformation Method) ranks among the simplest simulation techniques. This method produces random variables by applying the log-linear distribution function to each virtual parametric parameter value, given a specified sample size *N*, as follows

1. Generate random numbers from uniform distribution inside the interval [0, 1].
2. Apply the inverse transform method using the cumulative distribution function (CDF) to convert the generated uniform random variables into exponentially distributed variables. The procedure for generating exponential random variables via this technique proceeds as follows: $m(t) = \int_0^t \lambda(u)du$, where , $\lambda(t) = e^{\alpha+\beta t}$ $0 \le t \le \infty,\ \alpha, \beta > 0$, then

$$m(t) = \int_0^t e^{\alpha+\beta u}du = \frac{1}{\beta}\left[\left(e^{\alpha+\beta t}\right) - e^{\alpha}\right]$$
$$= \frac{e^{\alpha}}{\beta}\left(e^{\beta t} - 1\right)$$

Since $U = m(t)$ where $U \sim U(0,1),\ u = \frac{e^{\alpha}}{\beta}\left(e^{\beta t} - 1\right)$, then $e^{\beta t} = \frac{\beta\, u}{e^{\alpha}} + 1$ The following can be produced by calculating the natural logarithm of both sides: $\beta t = \ln\left(\frac{\beta\, u}{e^{\alpha}} + 1\right)$, this implies that $t = \frac{1}{\beta}\ln\left(\frac{\beta\, u}{e^{\alpha}} + 1\right)$

Consequently, the Inverse Transform Method is employed to generate a variety of random variables $t(i)$ according to the log-linear function, using a written MATLAB program.

Hence,

$$t(i) = \frac{1}{\beta}\ln\left(\frac{\beta}{e^{\alpha}}u(i) + 1\right) \text{ for } i = 1,2,\dots N \quad (16)$$

Where $U(i) \sim U(0,1)$

4.2.2 Thinning Method:

The thinning method is a simulation technique used to produce realizations of an NHPP. It begins by simulating an HPP with a constant rate that meets or exceeds the maximum of the target intensity function. The generated candidate points are then probabilistically filtered to ensure that the resulting process matches the intended time-varying intensity. The algorithm itself is simple and is applicable to all types of rate functions $\lambda(t)$. This approach is founded entirely on the following theorem, along with its proof presented by (Lewis and Shedler, 1979).

4.3Theorem (Thinning of a NHPP):

Examine a NHPP $\{N^*(t), t \ge 0\}$ expectation function and with rate function $\lambda^*(t)$:

$\mu^*(t) = \Lambda^*(t) - \Lambda^*(0)$. Let $T_1^*, T_2^*, \dots, T_{N*(T)}^*$ be the arrival times of $\{N^*(t)\}$ in the fixed interval $(0, T]$. Assume that for all t satisfying $0 \le t \le T$ it holds that $\lambda(t) \le \lambda^*(t)$.

For $i = 1,2,\dots, N^*(T)$, remove the point T_i^* with probability $1 - \lambda(T_i^*)/\lambda^*(T_i^*)$. Then the residual points form a NHPP $\{N(t), t \ge 0\}$ with rate function $\lambda(t)$ in the time interval $(0, T]$.

The universal algorithm for simulating a NHPP using the thinning method within the fixed interval $(0, T]$ is now outlined. The steps are as follows:

1 Generate the arrival times of a NHPP

$\{N(t)^*\}$ with rate function $\lambda^*(t)$ in the interval $(0,T]$. If the quantity of created points n^* equals 0, exit the algorithm → there are no actions of the procedure $\{N(t)\}$ in $(0,T]$.

2 Mean the made ordered arrival times by $t_1^*, t_2^*, \dots, t_{n^*}^*$. Set $i = 1$ and $k = 0$.

3 Create a variable u_i from the $U(0,1)$. If $u_i \le \lambda(t_i^*)/\lambda^*(t_i^*)$, set $k = k + 1$ and $t_k = t_i^*$.

4 Set $i = i + 1$. If $i \le n^*$, go to step 3.

5 Return $t_1, t_2, \dots, t_n$, where $n = k$, and also n.

The calculation of λ(t) is the primary cause of the algorithm's time inefficiency.

In the event that $\{N^*(t)\}$ is a HPP with rate function $\lambda^*(t) = \lambda^*$ and the minimum of $\lambda(t)$, represented by $\tilde{\lambda}$, is recognized, one can see that t_i^* is always known if $u \le \tilde{\lambda}/\lambda^*$. This might expedite the calculations, since one wouldn't always need to compute $\lambda(t)$.

4.4 Comparing Estimators:

When there are many methods used to estimate parameters, a comparison is conducted. In this research the measure RMSE is used. It is given by the following formula (Abbasi et al., 2006):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{Q} (\hat{\theta}_i - \theta)^2}{Q}},$$

$\hat{\theta}_i$: Denotes the parameter estimate calculated throughout the iteration i.

θ: Denotes the actual parameter value. Q : Denotes the number of iterations.

4.5 Computations Numerical and Discussion

When the proposed NHPP model with a log-linear rate function is applied to the arrival dataset, two simulation methods are used to generate random variables across different sample sizes $(15, 30, 45)$ and two sets of log-linear parameter estimates $(\hat{\alpha} = 7, 7.5$ and $\hat{\beta} = 0.001, 0.003)$. Subsequently, parameters are estimated using four approaches: classical methods (MLE and MoM) and advanced optimization techniques (PSO and Firefly Algorithms). To evaluate these estimating techniques' effectiveness, their results are compared based on the RMSE obtained from the simulation studies showed in Table 4. Among the 48 comparisons conducted, PSO generally outperformed the other techniques. Furthermore, Figure 7 displays the graphical representation of the log-linear function for actual passenger arrivals at EIA (January 1, 2021 - September 30, 2024) and simulation data-based using MLE, MoM, PSO, and FFA. It shows that PSO methods performs better than the other.

5.Conclusions

This paper proposed using an NHPP with a log-linear rate function to model passenger arrivals at EIA. The study used monthly passenger data from January 1, 2021, to September 30, 2024, to forecast future arrival volumes. Since the monthly number of arrivals at EIA varies independently from month to month, it is appropriately modeled as an NHPP; this was further supported by a homogeneity test confirming the data's nonhomogeneous nature. To estimate the key parameters of the log-linear rate function, four methods are employed: the classical approaches MLE and MoM, along with the intelligent optimization techniques PSO and FFA. An initial homogeneity test confirmed that the passenger arrival rate is nonhomogeneous and Cross Validation technique was applied to evaluate the model's predictive performance. Moreover, to assess the performance of these estimation methods, a simulation study was conducted using the real dataset, and the methods were compared by calculating their RMSE values. The numerical and graphical results demonstrate that across various parameter values and sample sizes, PSO consistently outperforms the other approaches, providing more accurate estimates than MLE, MoM, and FFA.

Acknowledgement

The authors express their gratitude and acknowledge the administration office of EIA for their support in providing data to enhance this research.

Table 1: The number of passengers arriving at EIA for the period from 1/1/2021 to 30/9/2024.

Number	Date	number of passengers(f_i)	Number	Date	number of passengers(f_i)
1	01-Jan-21	29697	24	01-Dec-22	72964
2	01-Feb-21	28716	25	01-Jan-23	71615
3	01-Mar-21	41308	26	01-Feb-23	65590
4	01-Apr-21	38055	27	01-Mar-23	84278
5	01-May-21	36620	28	01-Apr-23	84912
6	01-Jun-21	54956	29	01-May-23	84483
7	01-Jul-21	76963	30	01-Jun-23	92607
8	01-Aug-21	63970	31	01-Jul-23	121355
9	01-Sep-21	73766	32	01-Aug-23	105706
10	01-Oct-21	67265	33	01-Sep-23	102523
11	01-Nov-21	60597	34	01-Oct-23	86263
12	01-Dec-21	67176	35	01-Nov-23	76406
13	01-Jan-22	59326	36	01-Dec-23	89007
14	01-Feb-22	57921	37	01-Jan-24	84212
15	01-Mar-22	87764	38	01-Feb-24	75315
16	01-Apr-22	60245	39	01-Mar-24	84851
17	01-May-22	91120	40	01-Apr-24	89761
18	01-Jun-22	78526	41	01-May-24	75558
19	01-Jul-22	98349	42	01-Jun-24	101783
20	01-Aug-22	90215	43	01-Jul-24	118156
21	01-Sep-22	94029	44	01-Aug-24	99678
22	01-Oct-22	85051	45	01-Sep-24	98899
23	01-Nov-22	70795			

Table 2: The empirical intensity of passengers arriving at EIA from 1/1/2021 to 30/9/2024.

Month	Interval	Median Interval	Number arrivals	Intensity per day	Ln (Intensity per day)
01-Jan-21	[0,31)	15.5	29697	957.97	6.86
01-Feb-21	[31,59)	45	28716	1,025.57	6.93
01-Mar-21	[59,90)	74.5	41308	1,332.52	7.19
01-Apr-21	[90,120)	105	38055	1,268.50	7.15
01-May-21	[120,151)	135.5	36620	1,181.29	7.07
01-Jun-21	[151,181)	166	54956	1,831.87	7.51
01-Jul-21	[181,212)	196.5	76963	2,482.68	7.82
01-Aug-21	[212,243)	227.5	63970	2,063.55	7.63
01-Sep-21	[243,273)	258	73766	2,458.87	7.81
01-Oct-21	[273,304)	288.5	67265	2,169.84	7.68
01-Nov-21	[304,334)	319	60597	2,019.90	7.61
01-Dec-21	[334,365)	349.5	67176	2,166.97	7.68
01-Jan-22	[365,396)	380.5	59326	1,913.74	7.56
01-Feb-22	[396,424)	410	57921	2,068.61	7.63
01-Mar-22	[426,457)	441.5	87764	2,831.10	7.95
01-Apr-22	[457,487)	472	60245	2,008.17	7.60
01-May-22	[487,518)	502.5	91120	2,939.35	7.99
01-Jun-22	[518,548)	533	78526	2,617.53	7.87
01-Jul-22	[548,579)	563.5	98349	3,172.55	8.06
01-Aug-22	[579,610)	594.5	90215	2,910.16	7.98
01-Sep-22	[610,640)	625	94029	3,134.30	8.05
01-Oct-22	[640,671)	655.5	85051	2,743.58	7.92
01-Nov-22	[671,701)	686	70795	2,359.83	7.77
01-Dec-22	[701,732)	716.5	72964	2,353.68	7.76
01-Jan-23	[732,763)	747.5	71615	2,310.16	7.75
01-Feb-23	[763,791)	777	65590	2,342.50	7.76

01-Mar-23	[791,822)	836.5	84278	2,718.65	7.91
01-Apr-23	[822,852)	837	84912	2,830.40	7.95
01-May-23	[852,883)	867.5	84483	2,725.26	7.91
01-Jun-23	[883,913)	898	92607	3,086.90	8.03
01-Jul-23	[913,944)	928.5	121355	3,914.68	8.27
01-Aug-23	[944,975)	959.5	105706	3,409.87	8.13
01-Sep-23	[975,1005)	990	102523	3,417.43	8.14
01-Oct-23	[1005,1036)	1020.5	86263	2,782.68	7.93
01-Nov-23	[1036,1066)	1051	76406	2,546.87	7.84
01-Dec-23	[1066,1097)	1081.5	89007	2,871.19	7.96
01-Jan-24	[1097,1128)	1112.5	84212	2,716.52	7.91
01-Feb-24	[1128,1157)	1142.5	75315	2,597.07	7.86
01-Mar-24	[1157,1188)	1172.5	84851	2,737.13	7.91
01-Apr-24	[1188,1218)	1203	89761	2,992.03	8.00
01-May-24	[1218,1249)	1233.5	75558	2,437.35	7.80
01-Jun-24	[1249,1279)	1264	101783	3,392.77	8.13
01-Jul-24	[1279,1310)	1294.5	118156	3,811.48	8.25
01-Aug-24	[1310,1341)	1325.5	99678	3,215.42	8.08
01-Sep-24	[1341,1371)	1356	98899	3,296.63	8.10

Table3. Parameter Estimates for log-linear rate for passenger arrival data using different methods

Methods	Parameters estimate $\hat{\alpha}$	Parameters estimate $\hat{\beta}$
MLE	7.246	0.001
MoM	7.20	0.0008
PSO	7.832	0.0002
FFA	7.0	0.003

Table 4. Simulation-Based RMSE for Parameter Estimates Using MLE, MoM, PSO, and FFA

α	β	n	Methods	Simulated by thinning method then estimated parameter		Simulated by inverse method then estimated parameter	
				$RMSE(\hat{\alpha})$	$RMSE(\hat{\beta})$	$RMSE(\hat{\alpha})$	$RMSE(\hat{\beta})$
7	0.001	15	MLE	0.0005^*	0.0015	0.0000	0.0015
			MoM	3.2889	0.6835	3.2887	0.6839
			PSO	0.0020	0.0005^*	0.0000^*	0.0010
			FFA	0.2500	0.0000	0.2500	0.0000^*
		30	MLE	0.0015^*	0.0035	0.0015	0.0035
			MoM	3.2584	0.7330	3.2584	0.7331
			PSO	0.0050	0.0005	0.0000^*	0.0010
			FFA	0.2500	0.0000^*	0.1250	0.0000^*
		45	MLE	0.0045^*	0.0050	0.0045^*	0.0050
			MoM	3.2423	0.7611	3.2423	0.7611
			PSO	0.0085	0.0005	0.0050	0.0010^*
			FFA	0.2500	0.0000^*	0.1250	0.0010
7	0.003	15	MLE	0.0015^*	0.0040	0.0015	0.0040
			MoM	3.2912	0.6809	3.2912	0.6809
			PSO	0.0050	0.0010^*	0.0050	0.0010
			FFA	0.2500	0.0010	0.1250	0.0000
		30	MLE	0.0040^*	0.0095	0.0050^*	0.0100
			MoM	3.2634	0.7281	3.2627	0.7293
			PSO	0.0150	0.0010^*	0.0150	0.0010
			FFA	0.2500	0.0010	0.1250	0.0000^*
		45	MLE	0.0145^*	0.0155	0.0150^*	0.0160
			MoM	3.2495	0.7542	2.7693	0.7543
			PSO	0.0250	0.0010^*	0.0250	0.0010
			FFA	0.2500	0.0010	0.1250	0.0000^*
7.5	0.001	15	MLE	0.0000	0.0015	0.0000	0.0015
			MoM	3.5500	0.6724	3.5498	0.6727
			PSO	0.0000^*	0.0015	0.0000^*	0.0005^*
			FFA	0.5000	$0.0000*$	0.3750	0.0010
		30	MLE	0.0050	0.0035	0.0050	0.0035
			MoM	3.5212	0.7182	3.5212	0.7181
			PSO	0.0050^*	0.0030^*	0.0050^*	0.0005^*
			FFA	0.5000	0.0000	0.3750	0.0010
		45	MLE	0.0050	0.0050	0.0045^*	0.0050^*
			MoM	3.5057	0.7442	3.5057	0.7443
			PSO	0.0050^*	0.0050	0.0050	0.0005
			FFA	0.5000	0.0000^*	0.3750	0.0010
7.5	0.003	15	MLE	0.0050	0.0005	0.0000^*	0.0040
			MoM	3.5057	0.6722	3.5525	0.6695
			PSO	0.0050^*	0.0005^*	0.0050	0.0010
			FFA	0.5000	0.0010	0.3750	0.0000^*
		30	MLE	0.0050	0.0100	0.0050^*	0.0100
			MoM	3.5257	0.7137	3.5257	0.7136
			PSO	0.0000^*	0.0010^*	0.0150	0.0010
			FFA	0.5000	0.0010	0.3750	0.0000^*
		45	MLE	0.0150	0.0160	0.0140^*	0.0155
			MoM	3.5124	0.7379	3.5122	0.7381
			PSO	0.0000^*	0.0010^*	0.0250	0.0010
			FFA	0.5000	0.0010	0.3750	0.0000^*

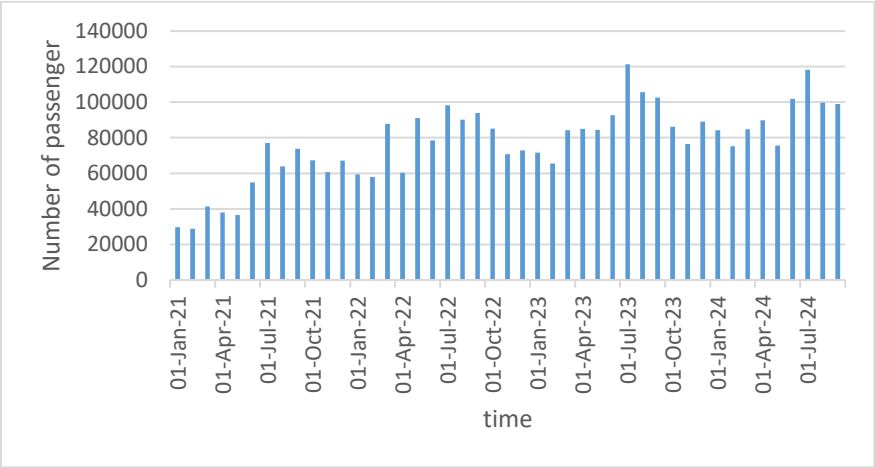


Figure 1. Passenger Arrivals at EIA from January 1, 2021, to September 30, 2024.

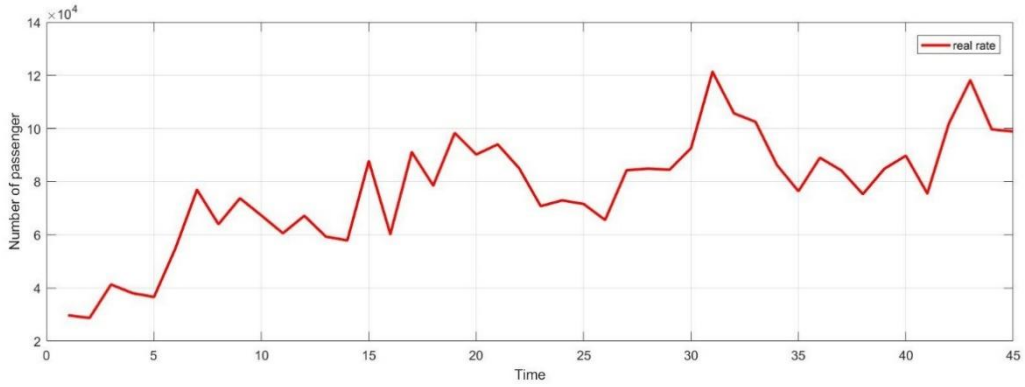


Figure 2. Intensity for Arrivals at EIA from January 1, 2021, to September 30, 2024.

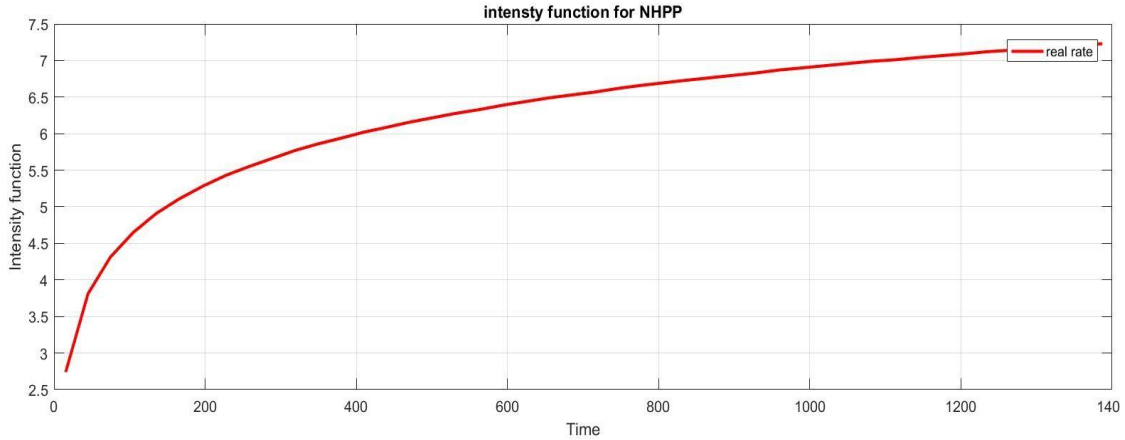


Figure 3. The intensity logarithmic of the arriving passengers from 1/1/2021 to 30/9/2024.

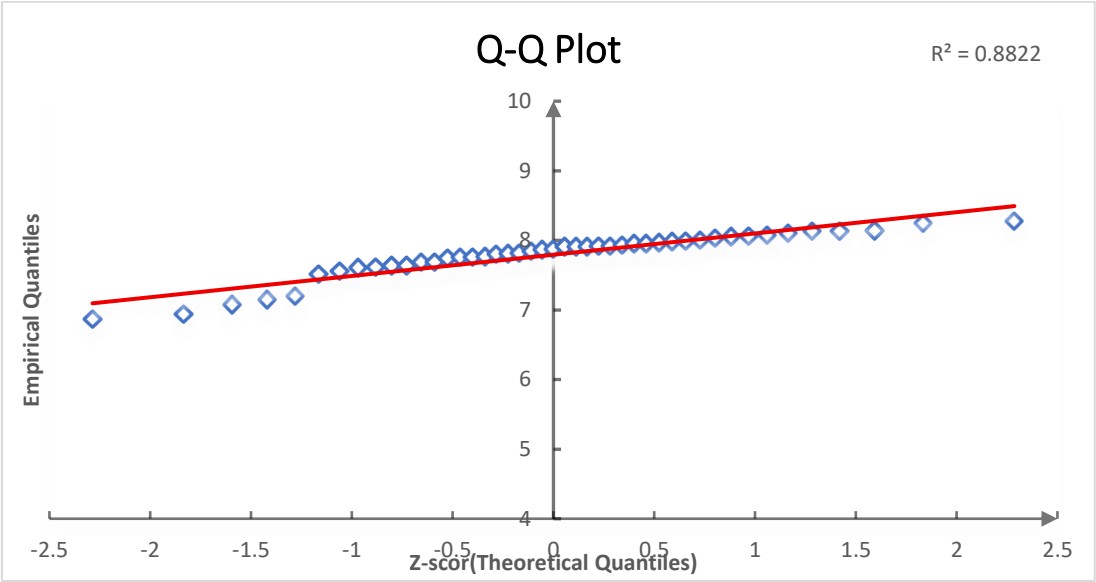
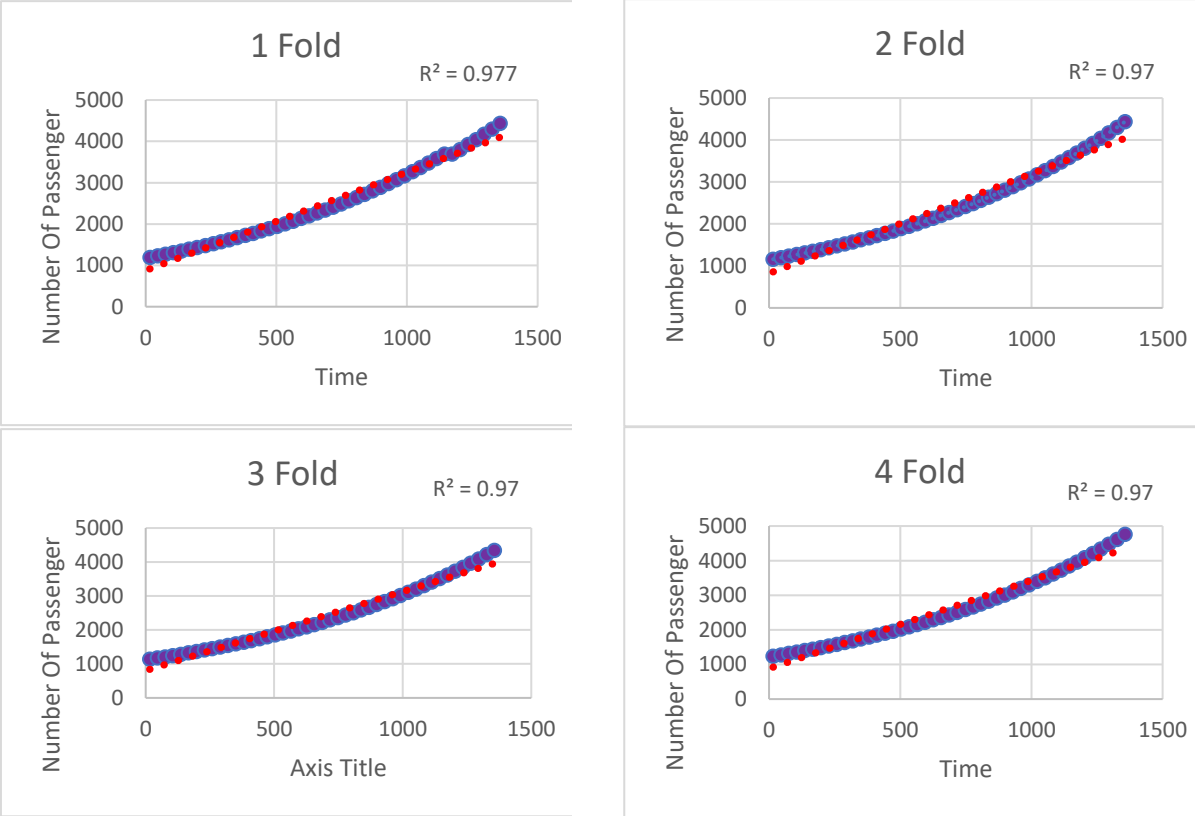


Figure 4. The Q-Q plot between empirical Quantiles and Z-Scor



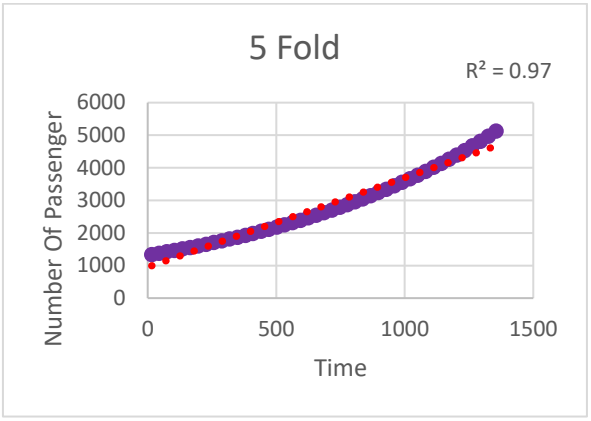


Figure 5. 5-fold Cross validation for the Number of Passenger

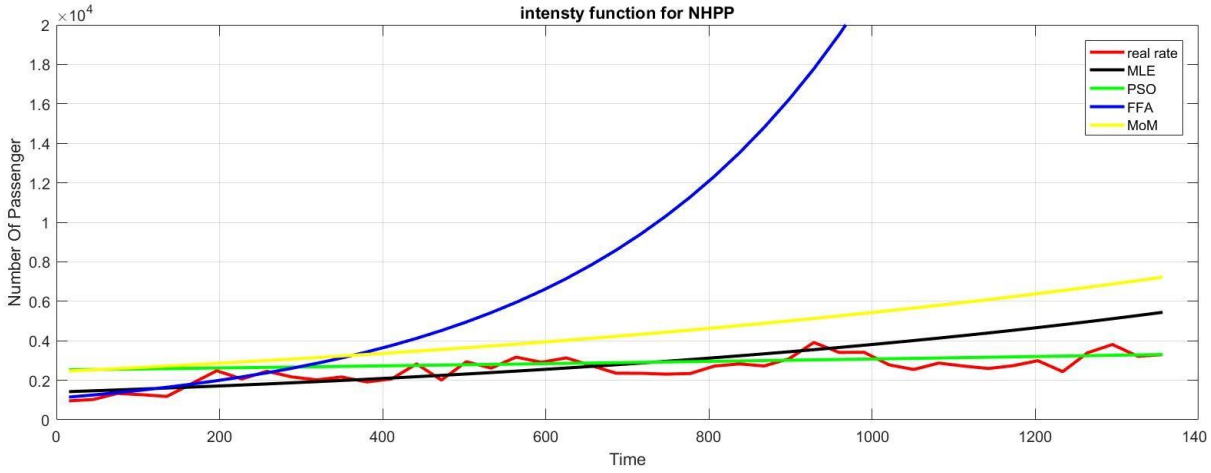


Figure 6.log-linear function for the real data compared to the Data-Based using MLE, MoM, PSO, and FFA

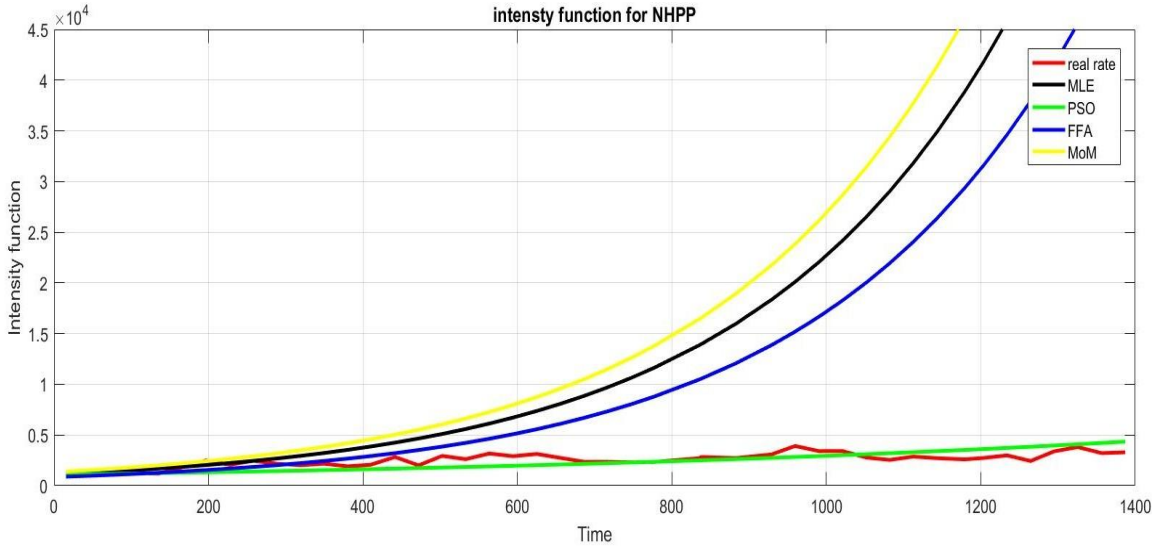


Figure 7. log-linear function for the real data compared to the Simulated Data-Based using MLE, MoM, PSO, and FFA

Reference:

ABBASI, B., JAHROMI, A. H. E., ARKAT, J. & HOSSEINKOUCHACK, M. 2006. Estimating the parameters of Weibull distribution using simulated annealing algorithm. *Applied Mathematics and Computation,* 183**,** 85-93.

ABDEL-BASSET, M., MOHAMED, R., HEZAM, I. M., SALLAM, K. M., ALSHAMRANI, A. M. & HAMEED, I. A. 2024. Artificial intelligence-based optimization techniques for optimal reactive power dispatch problem: a contemporary survey, experiments, and analysis. *Artificial Intelligence Review,* 58**,** 2.

ALI, M. S. & AL-ZUBAIDI, T. H. A. 2025. Using Some Linear Dynamic Systems with Bivariate Wavelets.

AMINZADEH, M. 2025. Copula-based Bayesian estimation of probabilities for two dependent non homogenous compound Poisson processes. *Communications in Statistics-Theory and Methods,* 54**,** 2320-2341.

ARORA, S. & SINGH, S. 2013. The firefly optimization algorithm: convergence analysis and parameter selection. *International Journal of Computer Applications,* 69.

BENING, J. E., DARKO, F., ANTERKYI, R., NARH, V. K., PREMPEH, N. Y. A., JAIN, N., KWANSA, A. L., KONTOR, E. K. & NTIM, M. 2025. Predictors of postpartum depression among new mothers in Kumasi, Ghana: A multicenter study using Bayesian analysis. *Women's Health,* 21**,** 17455057251343953.

BHASKAR, A., PONNURAJA, C., SRINIVASAN, R. & PADMANABAN, S. 2020. Distribution and growth rate of COVID-19 outbreak in Tamil Nadu: A log-linear regression approach. *Indian Journal of Public Health,* 64**,** 188-191.

BRATTON, D. & KENNEDY, J. Defining a standard for particle swarm optimization. 2007 IEEE swarm intelligence symposium, 2007. IEEE, 120-127.

BROWN, L. D. & ZHAO, L. H. 2002. A test for the Poisson distribution. *Sankhyā: The Indian Journal of Statistics, Series A***,** 611-625.

CHANG, G., GUIDA, W. C. & STILL, W. C. 1989. An internal-coordinate Monte Carlo method for searching conformational space. *Journal of the American Chemical Society,* 111**,** 4379-4386.

CHEN, S. & LÜ, X. 2024. Adaptive network traffic control with approximate dynamic programming based on a non-homogeneous Poisson demand model. *Transportmetrica B: Transport Dynamics,* 12**,** 2336029.

DASGUPTA, A. & DASGUPTA, A. 2008. Transformations. *Asymptotic Theory of Statistics and Probability***,** 49-61.

FISTER, I., FISTER JR, I., YANG, X.-S. & BREST, J. 2013. A comprehensive review of firefly algorithms. *Swarm and evolutionary computation,* 13**,** 34-46.

FRENKEL, I., GERTSBAKH, I. & KHVATSKIN, L. 2003. Parameter estimation and hypotheses testing for nonhomogeneous poisson process. *Transport Telecommun,* 4**,** 9-17.

GHILLI, D., RICCI, C. & ZANCO, G. 2024. A mean field game model for COVID-19 with human capital accumulation. *Economic Theory,* 77**,** 533-560.

GUO, H., LIN, K., YANG, K., MA, Z., CAO, M., HU, Y. & YAN, Y. 2022. Trends of cancer incidence among Chinese older adults from 2005 to 2016: A log-linear regression and age-period-cohort analysis. *Frontiers in Public Health,* 10**,** 1023276.

HASSEN, S. & ABDLRAZAQ, A. 2024. Contextual Deep Semantic Feature Driven Multi-Types Network Intrusion Detection System for IoT-Edge Networks. *Zanco Journal of Pure and Applied Sciences,* 36**,** 132-147.

HSU, H.-P. 2022. Solving the feeder assignment, component sequencing, and nozzle assignment problems for a multi-head gantry SMT machine using improved firefly algorithm and dynamic programming. *Advanced engineering informatics,* 52**,** 101583.

HUSSEIN, N. A. & HASSAN, R. 2023. Analysis of the contributory factors to accidents at signalized intersections using generalized estimating equation with negative binomial distribution. *Zanco Journal of Pure and Applied Sciences,* 35**,** 29-40.

ISMAEL, S. A. 2024. Predictors of mortality among critical COVID-19 patients admitted to the intensive care unit in the Sulaimani governorate in 2021, Iraq. *Zanco Journal of Medical Sciences (Zanco J Med Sci),* 28**,** 73-84.

JIANG, W. & SIMON, R. 2007. A comparison of bootstrap methods and an adjusted bootstrap approach for estimating the prediction error in microarray classification. *Statistics in medicine,* 26**,** 5320-5334.

KILINC, H. C. & YURTSEVER, A. 2022. Short-term streamflow forecasting using hybrid deep learning model based on grey wolf algorithm for hydrological time series. *Sustainability,* 14**,** 3352.

KWIECIEŃ, J. & FILIPOWICZ, B. 2012. Firefly algorithm in optimization of queueing systems. *Bulletin of the Polish Academy of Sciences. Technical Sciences,* 60**,** 363-368.

LEWIS, P. W. & SHEDLER, G. S. 1979. Simulation of nonhomogeneous Poisson processes by thinning. *Naval research logistics quarterly,* 26**,** 403-413.

LI, Q., GUO, F. & KIM, I. 2020. A non-parametric Bayesian change-point method for recurrent events. *Journal of Statistical Computation and Simulation,* 90**,** 2929-2948.

MURRAY, L. L. & BELLHOUSE, D. R. 2020. WF Sheppard's correspondence with Karl Pearson and the development of his tables and moment estimates. *Historia Mathematica,* 53**,** 108-117.

OFORI-ADDO, E. 2023. Modeling repairable system failure data using NHPP reliability growth mode.

OYEDELE, O. 2023. Determining the optimal number of folds to use in a K-fold cross-validation: A neural network classification experiment. *Research in mathematics,* 10**,** 2201015.

PARK, S. K. 2024. COMPARATIVE ANALYSIS ON THE PERFORMANCE OF NHPP-BASED SOFTWARE RELIABILITY MODEL FOLLOWING EXPONENTIAL LIFE DISTRIBUTION. *Journal of Theoretical and Applied Information Technology,*

102.
PATEL, R., NARMAWALA, S., MAHIDA, N., GUPTA, R., TANWAR, S. & SHAHINZADEH, H. ML-Based Optical Fibre Fault Detection in Smart Surveillance and Traffic Systems. 2024 15th International Conference on Information and Knowledge Technology (IKT), 2024. IEEE, 216-221.
ROSS, S. M. 2014. *Introduction to probability models*, Academic press.
SENCEROGLU, S., AYARI, M. A., REZAEI, T., FARESS, F., KHANDAKAR, A., CHOWDHURY, M. E. & JAWHAR, Z. H. 2022. Constructing an intelligent model based on support vector regression to simulate the solubility of drugs in polymeric media. *Pharmaceuticals,* 15**,** 1405.
SH, A., FATAH, K. S. & SULAIMAN, M. S. 2023. Estimating the Rate of Occurrence of Exponential Process Using Intelligence and Classical Methods with Application. *Palestine Journal of Mathematics,* 12.
SONG, M.-P. & GU, G.-C. Research on particle swarm optimization: a review. Proceedings of 2004 international conference on machine learning and cybernetics (IEEE Cat. No. 04EX826), 2004. IEEE, 2236-2241.
SUGIYANTO, G. & SANTI, M. Y. 2017. Road traffic accident cost using human capital method (Case study in Purbalingga, Central Java, Indonesia). *Jurnal Teknologi,* 79.
SZEMIS, J., PEDRUCO, P., LADSON, T. & NATHAN, R. Flike versus RMC-BestFit FFA Software-which one should I use? 2024 Hydrology and Water Resources Symposium (HWRS 2024), 2024. National Committee on Water Engineering, Engineers Australia Melbourne, 434-443.
VENKATARAMAN, S. & RUMPLER, R. 2025. Urban traffic flow estimation with noise measurements using log-linear regression. *Applied Acoustics,* 236**,** 110745.
VERONICA, K., ORAWO, L. A. O. & ISLAM, A. S. 2014. Likelihood Based Estimation of the Parameters of a Log-Linear Nonhomogeneous Poisson Process. *International Journal of Science and Research (IJSR),* 3**,** 200-204.
WANG, D., TAN, D. & LIU, L. 2018. Particle swarm optimization algorithm: an overview. *Soft computing,* 22**,** 387-408.
XIE, E., ZHU, L., ZHAO, L. & CHANG, L.-S. 1996. The human serotonin 5-HT2Creceptor: Complete cDNA, genomic structure, and alternatively spliced variant. *Genomics,* 35**,** 551-561.
YARMOHAMMADI, M., AFSHAR, A., MAHMOUDVAND, R. & NASIRI, P. 2022. Predicting intensity function of nonhomogeneous Poisson process. *Journal of Statistical Modelling: Theory and Applications,* 3**,** 39-50.
YU, T., SHAO, A., WU, H., SU, Z., SHEN, W., ZHOU, J., LIN, X., SHI, D., GRZYBOWSKI, A. & WU, J. 2025. A Systematic Review of Advances in AI-Assisted Analysis of Fundus Fluorescein Angiography (FFA) Images: From Detection to Report Generation. *Ophthalmology and therapy***,** 1-21.
ZHAO, J., WANG, Y., LU, H., LI, Z. & MA, X. 2021. Interference-based QoS and capacity analysis of VANETs for safety applications. *IEEE Transactions on Vehicular Technology,* 70**,** 2448-2464.